The UMBC High Performance Computing Facility

Matthias K. Gobbert

Department of Mathematics and Statistics
University of Maryland, Baltimore County (UMBC)

HPCF Governance Committee:
Matthias Gobbert (Math/Stat), Curtis Menyuk (CSEE), Marc Olano (CSEE), Lynn Sparling (Phys), Larrabee Strow (Phys), Ian Thorpe (Chem), Claire Welty (CUERE)

Funding from: NSF, NSA, NASA, IBM, UMBC, and more

Thanks to: HPCF RAs (Xuan Huang, Samuel Khuvis, Jonathan Graf, and others), DoIT staff (Schou, Philipp, Slettebak, Toothe, Souder, and more)!

UMBC High Performance Computing Facility (HPCF)

The community-based, interdisciplinary core facility for scientific computing and research on parallel algorithms at UMBC

- Initiated in 2008 with participation of over 20 faculty from more than 10 departments and research centers from all three colleges at UMBC
- Summer 2008: 35-node cluster (two dual-core AMD Opteron) with InfiniBand (dual data rate = DDR) and 14 TB central storage
- Replacement in Summer 2009: 86-node cluster (two quad-core Intel Nehalem) with new InfiniBand (quad data rate = QDR) and 160 TB central storage
- Original 32 nodes purchased with seed funding from UMBC and contributions from faculty; extension partially funded by NSF grants (in 2008: NSF MRI $200,000, SCREMS for Math & Stat $40,000) and contributions from faculty and UMBC
- System administration by Division of Information Technology; user support by HPCF RAs in collaboration with CIRC
- Governed by HPCF Governance Committee; point of contact: Matthias K. Gobbert, gobbert@umbc.edu, 410–455–2404
- All details, list of projects and publications, user info at `www.umbc.edu/hpcf`.

Purchases of Clusters in HPCF

- 2003 — kali (IBM): \$150,000
  33-node cluster (two single-core Intel Xeon)
  with Myrinet and 0.5 TB central storage
  funding: \$75k NSF SCREMS, \$75k UMBC cost-sharing, discount from IBM
- 2008 — hpc (IBM): \$270,000
  35-node cluster (two dual-core AMD Opteron)
  with InfiniBand (dual data rate = DDR) and 14 TB central storage
  funding: \$100k seed money UMBC, \$70k faculty, \$100,000 DoIT
- 2009 — tara (IBM): \$600,000
  86-node cluster (two quad-core Intel Nehalem)
  with new InfiniBand (quad data rate = QDR) and 160 TB central storage
  funding: \$360k faculty, \$240k NSF MRI and SCREMS
- 2013 — maya (Dell) — extension of tara: \$540,000
  72 nodes (19 hybrid CPU/GPU, 19 hybrid CPU/Phi, 34 CPU-only nodes;
  two eight-core Intel CPUs, two NVIDIA GPUs/Intel Phi per hybrid node)
  with InfiniBand (quad data rate = QDR) and over 750 TB central storage
  funding: NSF MRI \$300k, cost-sharing \$120k, existing \$260k
- Also in 2013: gift from NASA Goddard — essentially as 2009 tara!

Overview of the $1.5M 240-Node Cluster maya by Nodes

Heterogeneous cluster with three groups of nodes:

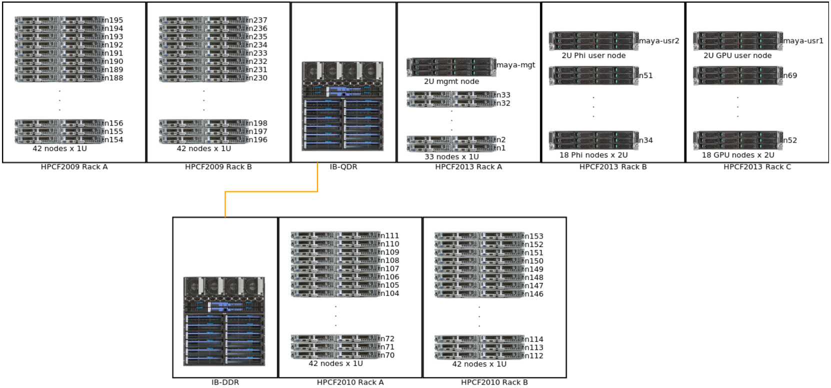
- **HPCF2009 = maya (2009) = tara (IBM):**
  84 nodes, each with two 2.6 GHz quad-core Intel Nehalem X5550 CPUs
- **HPCF2010 = maya (2010) = NASA Goddard gift (IBM):**
  84 nodes, each with two 2.8 GHz quad-core Intel Nehalem X5560 CPUs
- **HPCF2013 = maya (2013) = extension from Dell:**
  72 nodes, each with two 2.6 GHz eight-core Intel E5-2650v2 Ivy Bridge CPUs,
  34 CPU-only and 38 hybrid nodes:
  - 19 hybrid nodes with two NVIDIA K20 GPU
  - 19 hybrid nodes with two Intel Phi 5110P accelerators

Networks connecting all components:

- quad-data rate (QDR) InfiniBand interconnect for HPCF2009 and HPCF2013
- dual-data rate (DDR) InfiniBand interconnect for HPCF2010

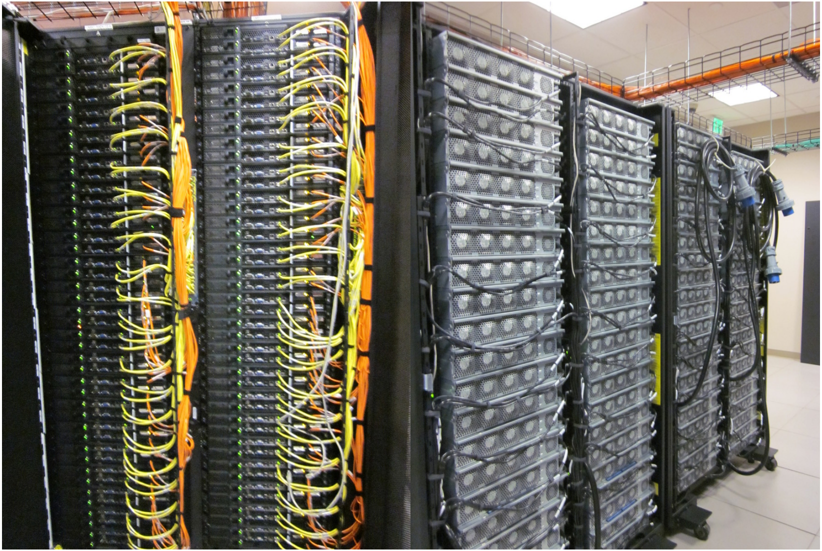
Central storage of over 750 TB connected by InfiniBand
For more information, see the webpage www.umbc.edu/hpcf

Maya: System Layout

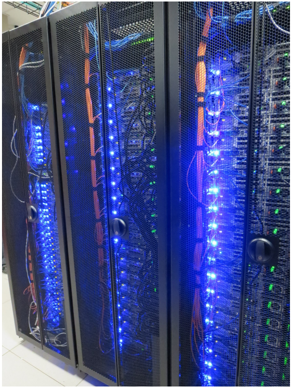


Maya HPCF2009: iDataPlex Rack and Detail

Maya HPCF2010: Racks Front and Back



Maya HPCF2013: Racks Front and Back



Maya HPCF2009, HPCF2013: Front and Back of the QDR InfiniBand Switch

Parabolic Nonlinear Three-Species Application Problems: CPU Only

Parabolic nonlinear three-species application problem (calcium wave), mesh resolution $N_x \times N_y \times N_z = 64 \times 64 \times 256$ [Huang and Gobbert, HPCF–2015–8, 2015]

Cluster, method	serial 1 core time	(1 node) all cores time (speedup)	32 node 1 core per node time (speedup)	32 node all cores time (speedup)
tara (2009), FEM	67:04:28	09:17:32 (7.22)	02:15:03 (29.80)	00:29:06 (138.29)
tara (2009), FVM	47:46:46	07:31:51 (6.34)	01:41:46 (28.17)	00:25:54 (110.68)
maya (2009), FVM	32:02:14	05:49:56 (5.49)	01:05:40 (29.27)	00:17:57 (107.09)
maya (2010), FVM	30:51:54	05:46:48 (5.34)	01:04:22 (28.77)	00:20:37 (89.83)
maya (2013), FVM	25:02:01	02:25:55 (10.29)	00:53:17 (28.19)	00:18:33 (80.97)*

- three-species application problem: system dimension 3,257,475
- hpc, tara, maya: ODE solver NDFk, $1 \leq k \leq 5$, 73,123 time steps, linear solver QMR (2009), BiCGSTAB (2010, 2013).
- asterisk indicates use of 8 processes per node to enable run
- ⇒ **yet better, more efficient, parallel code needed**
  ⇒ **need sophisticated numerical methods and parallelization that take full advantage of available state-of-the-art architectures!**

Weak Scalability of CICR Problem on maya 2009

- We double z-dimension of domain Ω and numerical mesh in z-direction along with doubling p, starting with $\Omega = (-6.4, 6.4) \times (-6.4, 6.4) \times (-32.0, 32.0)$ and $N_z = 4\,N_x = 4\,N_y$ for $p = 8$ on 1 node.
  Thus, the numerical mesh spacings Δx, Δy, Δz stay constant.
- Finite volume method with NDFk time-stepping to $t_{\text{fin}} = 100$.
- We also double the spacing of CRUs in the z-direction to keep the physiology comparable as possible.

(a) Total number of degrees of freedom for the scalar test problem

$N_x \times N_y$	$p = 8$	$p = 16$	$p = 32$	$p = 64$	$p = 128$	$p = 256$
16×16	56,355	112,710	225,420	45,0840	901,680	1,803,360
32×32	421,443	842,886	1,685,772	3,371,544	6,743,088	13,486,176
64×64	3,257,475	6,514,950	13,029,900	26,059,800	52,119,600	104,239,200
128×128	25,610,499	51,220,998	102,441,996	204,883,992	409,767,984	819,535,968

(a) Wall clock time T_p in HH:MM:SS

$N_x \times N_y$	$p = 8$	$p = 16$	$p = 32$	$p = 64$	$p = 128$	$p = 256$
16×16	00:00:15	00:00:14	00:00:13	00:00:13	00:00:14	00:00:14
32×32	00:02:26	00:02:21	00:02:14	00:02:06	00:02:01	00:01:56
64×64	00:27:18	00:26:35	00:25:08	00:23:50	00:22:54	00:21:35
128×128	05:03:11	04:50:14	04:38:24	04:34:29	04:20:49	04:17:25

(d) Wall clock time per time step in seconds

$N_x \times N_y$	$p = 8$	$p = 16$	$p = 32$	$p = 64$	$p = 128$	$p = 256$
16×16	0.01	0.01	0.01	0.01	0.01	0.01
32×32	0.04	0.04	0.05	0.05	0.05	0.05
64×64	0.40	0.41	0.41	0.41	0.43	0.43
128×128	3.67	3.69	3.78	3.97	4.00	4.20

Parabolic Nonlinear Three-Species Application Problems: CPU/GPU

Parabolic nonlinear 3-species application problem (calcium wave), wall clock time (speedup) against one 16-core node using the finite volume method with BiCG-STAB [Huang, Ph.D. Applied Mathematics, 2015]

nodes (GPU/node)	$32 \times 32 \times 129$	$64 \times 64 \times 256$	$128 \times 128 \times 512$
1 node (16 cores)	00:12:58	02:25:55	26:53:37
1 node (1 GPU)	00:16:47 (0.77)	01:44:13 (1.40)	15:32:34 (1.73)
1 node (2 GPUs)	00:11:47 (1.10)	00:59:53 (2.44)	08:18:18 (3.24)
2 nodes (16 cores)	00:07:24	01:10:26	13:56:38
2 nodes (1 GPU)	00:12:20 (1.05)	00:58:12 (2.51)	08:14:26 (3.26)
2 nodes (2 GPUs)	00:09:27 (1.37)	00:35:31 (4.11)	04:25:55 (6.07)
4 nodes (16 cores)	00:06:30	00:39:04	07:21:17
4 nodes (1 GPU)	00:09:46 (1.33)	00:34:36 (4.22)	04:20:56 (6.18)
4 nodes (2 GPUs)	00:08:25 (1.54)	00:24:21 (5.99)	02:28:00 (10.90)
8 nodes (16 cores)	00:07:26	00:25:46	03:54:47
8 nodes (1 GPU)	00:08:32 (1.52)	00:24:58 (5.84)	02:24:46 (11.15)
8 nodes (2 GPUs)	00:08:10 (1.59)	00:20:27 (7.14)	01:31:22 (17.66)
16 nodes (16 cores)	N/A	00:21:29	02:17:48
16 nodes (1 GPU)	00:08:19 (1.56)	00:20:45 (7.03)	01:30:06 (17.91)
16 nodes (2 GPUs)	00:08:11 (1.58)	00:19:57 (7.31)	01:06:17 (24.34)

CPU-only on $128 \times 128 \times 512$: 01:40:35 on 32 nodes, 01:23:57 on 64 nodes

Metrics of Success and Invitation

All details on HPCF webpage `www.umbc.edu/hpcf`:
list of projects, publications, detailed description, tutorials, and more

- Funding: equipment ≥\$540k (plus cost-share, plus gift), training ≥\$1.3M, plus research
- Since 2008, over 400 total distinct users. Currently, over 300 users with access to maya.
- During April 2015 alone, 71 active users from 28 research groups (very conservative figure!)
- Publications (2008–April 2015): 181 total, including 74 journal, 23 proceedings, and 17 theses.

Invitation:
To use HPCF, simply submit account request form,
then report outcomes (posters, publications, student theses)!